

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 17:25:07 ; Search time 669 Seconds  
(without alignments)  
9668.072 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagcgcgagcgagcgagga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1277	100.0	1277	9	US-09-844-864-1
2	799.4	62.6	814	9	US-09-844-864-18
3	752.2	58.9	809	9	US-09-844-864-22
4	273.6	21.4	305	9	US-09-844-864-21
5	260.6	20.4	375	9	US-09-844-864-25
6	98.8	7.7	123	9	US-09-844-864-19
7	98.8	7.7	123	9	US-09-844-864-23
8	72	5.6	105	9	US-09-844-864-20
9	70.4	5.5	105	9	US-09-844-864-24
10	55.2	4.3	1041	17	US-10-437-963-8597
11	54.2	4.2	1500	13	US-10-412-699B-1133
12	54.2	4.2	1500	16	US-10-374-780A-680
13	52.8	4.1	2664	17	US-10-437-963-52312
14	52.2	4.1	720	17	US-10-437-963-77205

15	51.6	4.0	987	16	US-10-084-846A-99	Sequence 99, Appl
16	51.6	4.0	59816	16	US-10-084-846A-1	Sequence 1, Appl
17	51.6	4.0	59816	16	US-10-084-846A-2	Sequence 2, Appl
18	51	4.0	1185	15	US-10-156-761-6879	Sequence 6879, Ap
19	51	4.0	9025608	15	US-10-156-761-1	Sequence 1, Appl
20	50.6	4.0	1190	17	US-10-437-963-36882	Sequence 36882, A
21	50.4	3.9	1388	17	US-10-767-701-14298	Sequence 14298, A
22	50.2	3.9	1673	13	US-10-412-699B-1140	Sequence 1140, Ap
23	50.2	3.9	1673	16	US-10-374-780A-687	Sequence 687, App
24	49.6	3.9	400	15	US-10-193-002-174	Sequence 174, App
25	49.6	3.9	400	15	US-10-084-843-179	Sequence 179, App
26	49.6	3.9	14800	9	US-09-954-456-1601	Sequence 1601, Ap
27	49.6	3.9	14800	15	US-10-269-909-61	Sequence 61, Appl
28	49.6	3.9	14800	17	US-10-717-597-183	Sequence 183, App
29	49.6	3.9	14835	15	US-10-240-965-113	Sequence 113, App
30	49	3.8	1092	17	US-10-767-701-9739	Sequence 9739, Ap
31	49	3.8	1839	12	US-09-758-759-150	Sequence 150, App
32	49	3.8	5235	17	US-10-627-370-3	Sequence 3, Appl
33	49	3.8	6792	17	US-10-627-370-1	Sequence 1, Appl
34	49	3.8	7175	16	US-10-375-253-13	Sequence 13, Appl
35	49	3.8	7177	14	US-10-033-026-7	Sequence 7, Appl
36	49	3.8	7362	16	US-10-375-253-11	Sequence 11, Appl
37	49	3.8	7364	9	US-09-954-456-1179	Sequence 1179, Ap
38	49	3.8	7364	14	US-10-033-026-5	Sequence 5, Appl
39	49	3.8	7376	14	US-10-033-026-3	Sequence 3, Appl
40	49	3.8	109519	12	US-09-758-759-1	Sequence 1, Appl
41	48.6	3.8	633	13	US-10-027-632-154723	Sequence 154723,
42	48.6	3.8	633	13	US-10-027-632-154724	Sequence 154724,
43	48.6	3.8	633	13	US-10-027-632-154725	Sequence 154725,
44	48.6	3.8	633	16	US-10-027-632-154723	Sequence 154723,
45	48.6	3.8	633	16	US-10-027-632-154724	Sequence 154724,

## ALIGNMENTS

## RESULT 1

US-09-844-864-1  
; Sequence 1, Application US/09844864  
; Patent No. US20020042926A1  
; GENERAL INFORMATION:  
; APPLICANT: Matzuk, Martin  
; APPLICANT: Ren, Yongsheng  
; APPLICANT: Wu, Xuemei  
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS  
; FILE REFERENCE: P01925US2 / 09807797 / OFA 99-48  
; CURRENT APPLICATION NUMBER: US/09/844,864  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/106,020  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: PCT/US99/25209  
; PRIOR FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1  
; LENGTH: 1277  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-844-864-1

Query Match	100.0%	Score 1277;	DB 9;	Length 1277;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1277;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAGCGCGGCGAGCGGGGAGCGCACCATGTTCCCGGAGCAGCGTTCCACCCCTGCGCG 60		
Db	1	AAGCGCGGCGAGCGGGGAGCGCACCATGTTCCCGGAGCAGCGTTCCACCCCTGCGCG 60		
Qy	61	CATCTTATCCGAGCGCCACCAAGCGGGATGGCTGGAGGTCGAGCCAGGGGCTGC 120		
Db	61	CATCTTATCCGAGCGCCACCAAGCGGGATGGCTGGAGGTCGAGCCAGGGGCTGC 120		
Qy	121	CGACCCGCGCCCCCTTCCTCCCGGCTACAGACAGCTCATGCGCGGAGTAGTC 180		





```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 123
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-19

Query Match          7.7%; Score 98.8; DB 9; Length 123;
Best Local Similarity 98.0%; Pred. No. 4e-19;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 795 TTTCAGTCTTTAGAGCAGAGAAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 854
Db 9 TGTTCAGTCTTTAGAGCAGAGAAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 68

Qy 855 GGAGAGCGCCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAGGT 896
Db 69 GGAGAGCGCCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAGGT 110

RESULT 7
US-09-844-864-23
; Sequence 23, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 123
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-23

Query Match          7.7%; Score 98.8; DB 9; Length 123;
Best Local Similarity 98.0%; Pred. No. 4e-19;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 795 TTTCAGTCTTTAGAGCAGAGAAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 854
Db 9 TGTTCAGTCTTTAGAGCAGAGAAGTACGGCTACTATCACTGCAAGACTGCAAAATCCGGTG 68

Qy 855 GGAGAGCGCCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAGGT 896
Db 69 GGAGAGCGCCTATGTGTGTGTGTGTCAGGCGCACCAAGTAAGGT 110

RESULT 8
US-09-844-864-20
; Sequence 20, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28

```

; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 105  
; TYPE: DNA  
; ORGANISM: mus musculus  
US-09-844-864-20

Query Match 5.6%; Score 72; DB 9; Length 105;  
Best Local Similarity 100.0%; Pred. No. 3.9e-11;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 958  
Db 19 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 78

QY 959 ATCACCTGTCAA 970  
Db 79 ATCACCTGTCAA 90

## RESULT 9

US-09-844-864-24  
; Sequence 24, Application US/09844864  
; Patent No. US20020042926A1  
; GENERAL INFORMATION:  
; APPLICANT: Matzuk, Martin  
; APPLICANT: Ren, Yongsheng  
; APPLICANT: Wu, Xuemei  
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS  
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48  
; CURRENT APPLICATION NUMBER: US/09/844,864  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/106,020  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: PCT/US99/25209  
; PRIOR FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24  
; LENGTH: 105  
; TYPE: DNA  
; ORGANISM: mus musculus  
US-09-844-864-24

Query Match 5.5%; Score 70.4; DB 9; Length 105;  
Best Local Similarity 98.6%; Pred. No. 1.2e-10;  
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 958  
Db 19 TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTTACAAACCTTACAGAGTGGAGGAC 78

QY 959 ATCACCTGTCAA 970  
Db 79 ATCACCTGTCAA 90

## RESULT 10

US-10-437-963-8597/c  
; Sequence 8597, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 8597  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15091C.1  
US-10-437-963-8597

Query Match 4.3%; Score 55.2; DB 17; Length 1041;  
Best Local Similarity 46.2%; Pred. No. 1.3e-05;  
Matches 258; Conservative 0; Mismatches 293; Indels 7; Gaps 2;

QY 89 GGGATGGCTGGAGTTCGGAGCCAGGGGCTGCGAGCCCGCGCCCCCTCTCTCTCTCCCCG 148  
Db 583 GGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524  
QY 149 GCTACAGACAGCTCATGCGCGCGGGAGTAGTGTACAGACAGCCACAGCGGGACAGCTCATGG 208  
Db 523 GCGGGGCGCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 465  
QY 209 CCTGTCTGTGCGGATGGGTCCCGCGTTCGGTTCAGCAGCGGTGACGCTGCGGTGACGTGA 268  
Db 464 GCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 405  
QY 269 ACCCGCGCGCGGACG-----CCTCGGTGAGTGTTCATCTCGGGCGCGCGACGCTGCAGC 322  
Db 404 CGCGCGGTCTCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345  
QY 323 CTGCAGGTCGCCAGCAGCCCGGAGCGCGCGGATCGGGTTCGTCAACCCCGTGGCCACG 382  
Db 344 GGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 285  
QY 383 CCGGCGCGGGAGATCCCGCGGATCTGTGAGACCGTGTAGCCCGTTCTCTCTCTGACCT 442  
Db 284 GGTGCGCGCGGGGGCGCGCGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225  
QY 443 TCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 502  
Db 224 GGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 165  
QY 503 AGGGAGAGCGGATCTCTCGGGAGCCCGGGAACCGGAGCCGAGAGAGTGGCGCGGAGGA 562  
Db 164 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105  
QY 563 AAGCGGTCCCGGAGCGGAGGAGGCGGATGTTTCTAGGCTGCGAGGCGAGGCGCGGT 622  
Db 104 CCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65  
QY 623 GGGAGCAGCAGCCAC 640  
Db 44 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27

## RESULT 11

US-10-412-699B-1133  
; Sequence 1133, Application US/10412699B  
; Publication No. US20040045049A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: Zhang, James  
; APPLICANT: Fromm, Michael E.  
; APPLICANT: Heard, Jacqueline E.  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Adam, Luc J.  
; APPLICANT: Broun, Pierre E.  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James S.  
; APPLICANT: Yu, Guo-Liang



[illegible]

## RESULT 13

```

US-10-437-52312/C
; Sequence 52312, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 52312
; LENGTH: 2664
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2664)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54620C.1
US-10-437-963-52312

```

[illegible]

## RESULT 14

```

US-10-437-963-77205/C
; Sequence 77205, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77205
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(720)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77125C.1
US-10-437-963-77205

```

Query Match 4.1%; Score 52.8; DB 17; Length 2664;

Db 366 AGTGGCGCGCGGCCACCGGAGCAGAGGAGGAGGAGCGGAGCGCGGAGCGCGGAGGCG 307  
Qy 608 CAGGCGAGGCGCGGTGGAGCAGCAGCCACACCGAGGACCGAACHGTGTGGCGGCGA 667  
Db 306 CCGCGCGGAGGCGCGGTGGAGAGCGCGAGCGCCGTTGGACGCCAGCAGCGCGCGGTGGG 247  
Qy 668 TGCAG 672  
Db 246 AGCGG 242

## RESULT 15

US-10-084-846A-99  
; Sequence 99, Application US/10084846A  
; Publication No. US2004006026A1  
; GENERAL INFORMATION:  
; APPLICANT: WEITNAUER, GABRIELE  
; APPLICANT: MUHLENWEG, AGNES  
; APPLICANT: TREPFER, AXEL  
; APPLICANT: BECHTHOLD, ANDREAS  
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES  
; FILE REFERENCE: 1974-005  
; CURRENT APPLICATION NUMBER: US/10/084,846A  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR APPLICATION NUMBER: PCT/EP01/09815  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: DE 101 09 166.4  
; PRIOR FILING DATE: 2001-02-25  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 99  
; LENGTH: 987  
; TYPE: DNA  
; ORGANISM: Streptomyces viridochromogenes  
; FEATURE:  
; OTHER INFORMATION: aviG4 dna: partial sequence of coding strand 1; nucleotide 1  
; OTHER INFORMATION: corresponds to nucleotide 45,341 of coding strand 1.  
US-10-084-846A-99

Query Match 4.0%; Score 51.6; DB 16; Length 987;  
Best Local Similarity 49.6%; Fred. No. 0.00015;  
Matches 132; Conservative 0; Mismatches 134; Indels 0; Gaps 0;  
Qy 80 CCAAAGCCGGGATGGCTGGAGGTTCCGGAGCCAGGGGCTGCCGACCCCGCGCCCTCCT 139  
Db 260 CCGCGTCCGACGCGCTGGCAGCACCTGGACACAGCGTCCGACCCGCGGCCCGGCT 319  
Qy 140 TCCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTGACAGCCACCCAGCGGCGAC 199  
Db 320 TCGACCAAGCTTCGGCCGGGACTTCTTCGGGTACTTGGCGGACGACCCCGACCTGCT 379  
Qy 200 AGCTCATGGCCCTGCTGCGGATGGTCCCGGTCCGTCAGTCAGCGCGTGGCGG 259  
Db 380 GGCTCTACACGGCGGATGAGCCAGGGCACCGGTGGATCCCGGCTTGGTCGGCGGC 439  
Qy 260 TGCAGGTGAACCCCGCGCGCGCTCGGTGCAAGTGTTCACCTCGGGCGCGCGACGCTGC 319  
Db 440 ACCAGGACTTCTCCGGCGTCCGACCGGTGGTGGAGCTCGGGGGCGGGACCGGAGCTGC 499  
Qy 320 AGCTGAGGGTGGAGCCAGCCCC 345  
Db 500 TGGCGCGGTGCTGCGCGCGCACCC 525

Search completed: September 24, 2004, 20:14:47  
Job time : 680 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 17:12:32 ; Search time 3653 Seconds  
(without alignments)  
10439.092 Million cell updates/sec

Title: US-09-830-810A-1  
Perfect score: 1277  
Sequence: 1 aagcgggcgagggcgggga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estmu:\*\*

4: em\_estnu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_estl:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rod:\*\*

26: em\_gss\_pg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	643	50.4	675	10	BB641267
2	642.4	50.3	666	10	BB264222
3	591.4	46.3	615	10	EG071693
4	582.4	45.6	595	12	EG084538

C	5	570.4	44.7	594	10	BF471866	BF471866 UI-M-BH3-
C	6	497.2	38.9	525	10	BB704019	BB704019 BB704019
C	7	473.8	37.1	499	9	AI854700	AI854700 UI-M-BH0-
C	8	448.4	35.1	491	10	BB703259	BB703259 BB703259
C	9	413.8	32.4	491	10	BB704648	BB704648 BB704648
C	10	405.8	31.8	436	9	AU044294	AU044294 AU044294
C	11	398.2	31.2	425	10	BB705931	BB705931 BB705931
C	12	396.4	31.0	590	9	AU023153	AU023153 AU023153
C	13	393	30.8	450	10	BE946858	BE946858 UI-M-BH3-
C	14	392.2	30.7	419	10	BB704449	BB704449 BB704449
C	15	367.4	28.8	521	10	BB703869	BB703869 BB703869
C	16	367	28.7	427	10	BB706957	BB706957 BB706957
C	17	364.2	28.5	434	10	BB699732	BB699732 BB699732
C	18	347.2	27.2	391	10	BB704146	BB704146 BB704146
C	19	333.8	26.1	400	10	BB700620	BB700620 BB700620
C	20	319	25.0	905	28	BZ114495	BZ114495 CH230-421
C	21	316.2	24.8	441	10	BB704848	BB704848 BB704848
C	22	282.2	22.1	301	13	EX637052	EX637052 EX637052
C	23	264.6	20.7	315	10	BB248342	BB248342 BB248342
C	24	227.4	17.8	642	13	BY719704	BY719704 BY719704
C	25	224.8	17.6	654	28	AZ225053	AZ225053 RPCI-23-6
C	26	211.6	16.6	251	9	AV359580	AV359580 AV359580
C	27	208.6	16.3	251	9	AV358045	AV358045 AV358045
C	28	203.2	15.9	249	9	AV357766	AV357766 AV357766
C	29	203	15.9	290	10	BB063113	BB063113 BB063113
C	30	201.8	15.8	539	9	AL904756	AL904756 AL904756
C	31	200.8	15.7	909	13	EX755136	EX755136 BX755136
C	32	200	15.7	286	10	BB060381	BB060381 BB060381
C	33	199.8	15.6	246	10	BB060134	BB060134 BB060134
C	34	198	15.5	539	9	AL904892	AL904892 AL904892
C	35	197.8	15.5	642	12	BM861479	BM861479 fy46d03.x
C	36	197	15.4	643	12	BU098465	BU098465 BU098465
C	37	196.8	15.4	526	9	AL904839	AL904839 AL904839
C	38	196.8	15.4	534	9	AL904820	AL904820 AL904820
C	39	196.8	15.4	559	9	AL904774	AL904774 AL904774
C	40	196.8	15.4	575	9	AL904855	AL904855 AL904855
C	41	196.4	15.4	576	9	AL904863	AL904863 AL904863
C	42	196.4	15.4	578	9	AL904817	AL904817 AL904817
C	43	196.4	15.4	579	9	AL904815	AL904815 AL904815
C	44	196.4	15.4	579	9	AL904816	AL904816 AL904816
C	45	196.4	15.4	604	9	AL904765	AL904765 AL904765

## ALIGNMENTS

RESULT 1	BB641267	675 bp	linear	EST 26-OCT-2001
LOCUS	BB641267	RIKEN full-length enriched, 10 days neonate cortex Mus		
DEFINITION	musculus cDNA clone A830014H23 5', mRNA sequence.			
ACCESSION	BB641267			
VERSION	BB641267.1	GI:16476392		
KEYWORDS	EST.			
SOURCE	Mus musculus			
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 675)			
AUTHORS	Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, P., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001)			
TITLE	Unpublished (2001)			
JOURNAL	Contact: Yoshihide Hayaishizaki			
COMMENT	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute			
	The Institute of Physical and Chemical Research (RIKEN)			
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan			



**THE UNIVERSITY OF CHICAGO**

**Table 1. Characteristics of the study population**

FEATURES  
source









sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.  
e mouse tissues.

FEATURES  
source

Location/Qualifiers  
1..491  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420449J15"  
/sex="female"  
/tissue\_type="in vitro fertilized eggs"  
/dev\_stage="egg"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, in vitro fertilized eggs"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
GAGAGAGAGAGATCCCAAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid KS(+) after bulk excision from Lambda PLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

## ORIGIN

Query Match 35.1%; Score 448.4; DB 10; Length 491;  
Best Local Similarity 98.2%; Pred. No. 8.8e-58;  
Matches 485; Conservative 0; Mismatches 6; Indels 3; Gaps 3;  
QY 763 CAAGCAGGAGCAGGAGCGCTCGTTCCAGTTCTTAGACGAGAGTACGGC 822  
Db 1 CAAGCAGGAGCAGGAGCGCTCGTTCCAGTTCTTAGACGAGAGTACGGC 60  
QY 823 TACTATCACTCAAGGACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGGTGTGCAG 882  
Db 61 TACTATCACTGCAA-GACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGGTGTGCAG 119  
QY 883 GGCACCACTAGGTTTACTTCAACAGTTCTGCGAGTGTGTGAGAAATCTCAACCC 942  
Db 120 GGCACCACTAGGTTG-TACTTAAACAGTTCTGCGAGTGTGTGAGAAATCTCAACCC 178  
QY 943 TTACAGAGTGGAGGACATCACTGTCAAGTTGTAAAGCACTAGATGTGCTGCCAGT 1002  
Db 179 TTACAGAGTGGAGGACATCACTGTCAAGTTGTAAAGCACTAGATGTGCTGCCAGT 238  
QY 1003 CAGATTTCCGACGTGGACCTTAAACGCCGCCCATCGGAAGACTTGTGTGGAGATGCAA 1062  
Db 239 CAGATTTCCGACGTGGACCTTAAACGCCGCCCATCGGAAGACTTGTGTGGAGATGCAA 298  
QY 1063 GGACAAACGCCGTCTCGGACAGACCTTACAGTTTCAATACATCATTTAGTAGAGTC 1122  
Db 299 GGACAAACGCCGTCTCGGACAGACCTTACAGTTTCAATACATCATTTAGTAGAGTC 358  
QY 1123 GAAACGTTTCTGTAGATGGGCGCTAATGGAATGACAGAGCTTCTCCCTCTTCA 1182  
Db 359 GAAACGTTTCTGTAGATGGGCGCTAATGGAATGACAGAGCTTCTCCCTCTTCA 418

QY 1183 CCTCTCCCTTCCAAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGAA 1242  
Db |||||||

419 CCTCTCCCTTCCAAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGAA 477  
Db |||||||

QY 1243 TAAAGGTATTGCA 1256  
Db |||||||

478 TAAAGGTATTGCA 491  
Db |||||||

## RESULT 9

BB704648

LOCUS

DEFINITION

BB704648 RIKEN full-length enriched, in vitro fertilized eggs Mus

musculus cDNA clone 7420466L07 3', mRNA sequence.

ACCESSION

BB704648

VERSION

BB704648.1 GI:16053483

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 491)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,

Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,

Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,

Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akai,H.,

Tanaka,T., Tomaru,A., Toya,I., Watahiki,A., Yasunishi,A.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

JOURNAL

COMMENT

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,

Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) forfurther details.  
e mouse tissues.

## FEATURES

source

Location/Qualifiers

1..491

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CS7BL/6J"

/db\_xref="taxon:10090"

/clone="7420466L07"

/sex="female"

/tissue\_type="in vitro fertilized eggs"

/dev\_stage="egg"







```

Best Local Similarity 86.1%; Pred. No. 5.3e-50;
Matches 464; Conservative 0; Mismatches 71; Indels 4; Gaps 3;

QY 713 TGGCTCAGGACCCCGGTGATTCGATGCGCTCCAGACCGGCTCCCGCAAGACGCG 772
Db 542 TGGNCTNAGACCCCGGTATTCGATGCGCTCCAGACCGGCTCCCGCAAGACGCG 483
QY 773 AGCAGGACAAAGGAGCGCTGGTTCAGATTCCTTAGAGCAGCAAGTACGG--CTACTATCA 830
Db 482 AGNACAGGAGCGCGCTGCNTTTCAGAGCTCTTAGAGCAGCAAGTACGNCCTACTATCA 423
QY 831 CTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCTATGTG-TGGTGTGTGAGGCGCACCA 889
Db 422 CTGCAAGGAGCTGCAAAATCCGNTGGGAGAGCNCCTATGTGTGTGTGAGGCGCACCC 363
QY 890 GTAAGGTGTACTTCAACAGTCTCGCGAGTGTGTGAGAACTCTTACACCTTACAGA 949
Db 362 AGTAAAGTGTACTTCAACAGTCTCTGCGAGTGTGTGAGAACTCTTACACCTTACAGA 303
QY 950 GTGAGGACATCACCTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAGATTT 1009
Db 302 GTGAGGACATCACCTGTCAAAAGTTGTAAGAACTAGATGTGCTGCCAGTCAGATTT 243
QY 1010 CGCAGTGTGACCTTAAACGCCGCCATCGGCAAGACTGTGTGGAGATGCAAGGACAAA 1069
Db 242 CGCAGTGTGACCTTAAACGCCGCCATCGGCAAGACTGTGTGGAGATGCAAGGACAAA 183
QY 1070 CGCTGTGCTGCGACAGCAGCTTCAAGTCTCAATATACATCATTTAGTAGAGTCGAAAACG 1129
Db 182 CGCTGTGCTGCGACAGCAGCTTCAAGTCTCAATATACATCATTTAGTAGAGTCGAAAACG 123
QY 1130 TTTCTGTAGATGGGCTAAATGAATGAGCAAGTACGCTTCTCCCTCTTCACTCTTC 1189
Db 122 TTTCTGTAGATGGGCTAAATGAATGAGCAAGTACGCTTCTCCCTCTTCACTCTTC 63
QY 1190 CTTTCAAAATCTTATGACAGACAGTGTACTTGGATATAAAGCTGTGMAATAAAG 1248
Db 62 CTTTCAAAATCTTATGACAGACAGTGTACTTGGATATAAAGCTGTGMAATAAAG 5

RESULT 13
BE946858
LOCUS
DEFINITION
UT-M-BH3-awu-b-08-0-UI.s1 NIH BMAP M S4 Mus musculus cDNA clone
UT-M-BH3-awu-b-08-0-UI 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: m85@mail.nih.gov
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements

```

Seq primer: M13 Forward  
POLYA=NO.

#### FEATURES

source  
1..450  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-awu-b-08-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/dev\_stage="27-32 days"  
/clone\_lib="NIH BMAP M S4"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
NIH BMAP M S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged  
normalized libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH BMAP M S4,  
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,  
NIH BMAP M S2, NIH BMAP M S1. The subtracted library  
(NIH BMAP M S4) was constructed as follows: PCR amplified  
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and  
NIH BMAP M S3.1 clones from which 3' ESTs had been derived  
was used as a driver in a hybridization with a pool of  
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1  
libraries in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library)  
was purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (Life Technologies) to generate the  
NIH BMAP M S4 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)  
TAG\_SEQ=None found"

#### ORIGIN

Query Match 30.8%; Score 393; DB 10; Length 450;  
Best Local Similarity 99.7%; Pred. No. 1.9e-49;  
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 GCGCGGAGCGCGGAGCGACCCATGTTCCCGGAGCAGCATTCACCCCTGCCGCA 62  
Db 57 GCGCGGAGCGCGGAGCGACCCATGTTCCCGGAGCAGCATTCACCCCTGCCGCA 116  
QY 63 TCCTTATCCGAGGCCACCAAGCCGGGATGCTGGAGTTCGGAGCCAGGGCTGCCG 122  
Db 117 TCCTTATCCGAGGCCACCAAGCCGGGATGCTGGAGTTCGGAGCCAGGGCTGCCG 176  
QY 123 ACCCGCGCCCTCTCTTCTCCCGGCTACACAGCTCATCGCGCGAGTACGTCGA 182  
Db 177 ACCCGCGCCCTCTCTTCTCCCGGCTACACAGCTCATCGCGCGAGTACGTCGA 236  
QY 183 CAGCCACAGCGGCGACAGCTCATGSCCTGTGTCTCGCGATGGGTCCCGGTCTGTCAG 242  
Db 237 CAGCCACAGCGGCGACAGCTCATGSCCTGTGTCTCGCGATGGGTCCCGGTCTGTCAG 296  
QY 243 CAGCCGTGACGCTGCGGTGACAGTGAAACCGCGCGGAGCGCTCGGTGAGTGTCACT 302  
Db 297 CAGCCGTGACGCTGCGGTGACAGTGAAACCGCGCGGAGCGCTCGGTGAGTGTCACT 356  
QY 303 CCGGCGCCCGCAGCTGACGCTGCGGGTCCGAGCCAGCCAGCCGCGGATCGGGTTC 362  
Db 357 CCGGCGCCCGCAGCTGACGCTGCGGGTCCGAGCCAGCCAGCCGCGGATCGGGTTC 416  
QY 363 CTGTCAACCCCGTGGCCACGCGCGGCGCGGGAGA 396  
Db 417 CTGTCAACCCCGTGGCCACGCGCGGCGCGGGAGA 450

RESULT 14  
 BB704449  
 LOCUS  
 DEFINITION BB704449 RIKEN full-length enriched, in vitro fertilized eggs Mus  
 musculus cDNA clone 742046A04 3', mRNA sequence.  
 ACCESSION BB704449.1 GI:16053284  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Mus musculus (house mouse)

REFERENCE  
 AUTHORS  
 1 (bases 1 to 419)  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

TITLE  
 JOURNAL  
 COMMENT  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.  
 e mouse tissues.

PEATURES  
 Location/Qualifiers  
 1. 419  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="742046A04"  
 /sex="female"  
 /tissue\_type="in vitro fertilized eggs"  
 /dev\_stages="egg"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, in vitro fertilized eggs"  
 /note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'  
 GAGAGAGAGATTCGAGTTAAATTAATCCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI"

## ORIGIN

Query Match 30.7%; Score 392.2; DB 10; Length 419;  
 Best Local Similarity 98.8%; Pred. No. 2.5e-49;  
 Matches 416; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
 QY 834 CAAGAGCTGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGCGAGGCGACCACTAA 893  
 Db 1 CAAGAGCTGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGCGAGGCGACCACTAA 60  
 QY 894 GGCTGTACTTCAAAACAGTTCTGCCGAGTGTCTGAGAAATCTCAACCCCTTACAGAGTGG 953  
 Db 61 GGT-TTACTTCAAAACAGTTCTGCCGAGTGTCTGAGAAATCTCAACCCCTTACAGAGTGG 119  
 QY 954 AGGACATCACCCTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCCGAGTGCAGATTTGCC 1013  
 Db 120 AGGACATCACCCTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCCGAGTGCAGATTTGCC 179  
 QY 1014 ACCTGGAGCCCTAAAGCGCCCATCGGAGAGTGTGTGGGAGATGCGAGGACCAACGCC 1073  
 Db 180 ACCTGGAGCCCTAAAGCGCCCATCGGAGAGTGTGTGGGAGATGCGAGGACCAACGCC 239  
 QY 1074 TGCTCTGCGACAGCACCTTTCAGCTTCAAATACATCATTTAGTGAGAGTCGAAAAGCTTTC 1133  
 Db 240 TGCTCTGCGACAGCACCTTTCAGCTTCAAATACATCATTTAGTGAGAGTCGAAAAGCTTTC 299  
 QY 1134 TGCTAGATGGGGCTAAATGGAATGGAACAGTGTGCTTCCCTCTTCCCTCTTCCCTT 1193  
 Db 300 TGCTAGATGGGGCTAAATGGAATGGAACAGTGTGCTTCCCTCTTCCCTCTTCCCTT 359  
 QY 1194 TCCAAATCTTCATGACAGACAGTGTTCCTTCTTGGATATTAAGCCTGTGATTAAGGTATT 1253  
 Db 360 TCCAAATCTTCATGACAGACAGTGT-TACTTGGATATTAAGCCTGTGATTAAGGTATT 418  
 QY 1254 G 1254  
 Db 419 G 419

## RESULT 15

BB703869

LOCUS

DEFINITION

BB703869 RIKEN full-length enriched, in vitro fertilized eggs Mus

musculus cDNA clone 742045C21 3', mRNA sequence.

ACCESSION

BB703869

VERSION

BB703869.1 GI:16052704

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 521)

AUTHORS

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.



Blank

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 15:24:32 ; Search time 5112 Seconds  
(without alignments)  
10827.280 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aaggcgaggcgagcgaggga.....acaaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1245.4	97.5	1260	10	AY191415	AY191415 Mus muscu
2	562.2	75.3	1280	10	AY283175	AY283175 Rattus no
3	799.4	62.6	3987	10	AY193889	AY193889 Mus muscu
4	799.4	62.6	147464	2	AC122733	AC122733 Mus muscu
5	799.4	62.6	200535	10	AL671880	AL671880 Mouse DNA
6	689.8	54.0	168073	2	AC107686	AC107686 Mus muscu
7	628	49.2	235390	2	AC126519	AC126519 Rattus no
8	628	49.2	237695	2	AC125993	AC125993 Rattus no
9	628	49.2	262139	2	AC127083	AC127083 Rattus no
10	360	28.2	212848	2	AC108848	AC108848 Mus muscu
11	352	27.6	1275	9	AY191416	AY191416 Homo sapi
12	299.8	23.5	212848	2	AC108848	AC108848 Mus muscu
13	205.8	16.1	1052	5	AY283176	AY283176 Xenopus l
14	194.8	15.3	963	5	AY283177	AY283177 Takifugu
15	194.6	15.2	1084	5	AY283178	AY283178 Danio rer
16	147.6	11.6	138890	9	AC007970	AC007970 Homo sapi
17	147.6	11.6	151961	2	AC037443	AC037443 Homo sapi
18	147.6	11.6	204268	2	AC092163	AC092163 Homo sapi
19	134.6	10.5	4138	9	AY193890	AY193890 Homo sapi
20	134.6	10.5	187578	9	AC096952	AC096952 Homo sapi
21	90.6	7.1	187278	2	BX571888	BX571888 Danio rer
22	90.6	7.1	249791	5	BX537133	BX537133 Zebrafish
23	73.2	5.7	694	9	HS4335296	AJ335296 Homo sapi
24	72.2	5.7	125020	9	AF429315	AF429315 Homo sapi
25	64.8	5.1	144000	9	AC123764	AC123764 Homo sapi
26	63.2	4.9	125020	9	AF429315	AF429315 Homo sapi
27	62.6	4.9	1393	11	PM11H12G	AL684264 Penicilli
28	62	4.9	136551	2	AC048354	AC048354 Homo sapi
29	58.6	4.6	201858	2	AC115880	AC115880 Mus muscu
30	58.4	4.6	248550	1	SC0939120	AL939120 Streptomy
31	58.2	4.6	94058	2	AC141024	AC141024 Rattus no
32	58	4.5	92310	2	AC108380	AC108380 Pan trogl
33	57.8	4.5	2000	6	AX655393	AX655393 Sequence
34	57.2	4.5	143406	10	AC092857	AC092857 Rattus no
35	57.2	4.5	230329	2	AC142070	AC142070 Rattus no
36	56.6	4.4	245210	2	AC137771	AC137771 Homo sapi
37	56.2	4.4	257784	2	AC118627	AC118627 Mus muscu
38	56	4.4	247641	5	EX470223	EX470223 Zebrafish
39	56	4.4	248640	2	EX569785	EX569785 Danio rer
40	56	4.4	259098	2	EX548044	EX548044 Danio rer
41	55.6	4.4	991	11	PM12H12B	AL684455 Penicilli
42	55.6	4.4	151349	10	AL355176	AL355176 Mouse DNA
43	55	4.3	745	9	HS4335067	AJ335067 Homo sapi
44	54.6	4.3	183305	2	EX640474	EX640474 Danio rer
45	54.4	4.3	976	11	PM12D6G	AL684370 Penicilli

ALIGNMENTS

RESULT 1  
AY191415  
LOCUS AY191415 1260 bp mRNA linear ROD 12-MAR-2003  
DEFINITION Mus musculus zygote arrest 1 (Zarl) mRNA, complete cds.  
ACCESSION AY191415  
VERSION AY191415.1 GI:27808689  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1260)  
AUTHORS Wu,X., Viveiros,M.M., Eppig,J.J., Bai,Y., Fitzpatrick,S.L. and Matzuk,M.M.  
TITLE Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for

Pred. No. is the number of results predicted by chance to have a







JOURNAL	Submitted (09-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 3987)
AUTHORS	Wu, X. and Matzuk, M.M.
TITLE	Direct Submission
JOURNAL	Submitted (21-JAN-2003) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REMARK	Sequence update by submitter
COMMENT	On Jan 21, 2003 this sequence version replaced gi:27808693.
FEATURES	Location/Qualifiers
source	1..3987
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/strain="129/SvEv"
	/db_xref="taxon:10090"
gene	1..3987
	/gene="Zarl1"
mRNA	join(1..801,2574..2666,2745..2819,3697..3987)
	/gene="Zarl1"
	/product="zygote arrest 1"
exon	1..801
	/gene="Zarl1"
	/number=1
CDS	join(28..801,2574..2666,2745..2819,3697..3984)
	/product="zygote arrest 1"
	/protein_id="AA024708.1"
	/db_xref="GI:27808694"
	/translation="MFPASTFHPCHPYPQATKAGDGNRFQAGRCRAPPSPFLPGYRQ LMAEYVDSHQRAQLMALLSRMGRSVSRDAAVONPRDASVCSLGRRTLOPAGC RASPABVDSQCPGHAGRAGRSQWTVAPESVTFQGLSSLVACGRPTPKGEG SPASGSTRPEREVNARKAVQPRSEGDVQAQAGHEQOPPEPDNRNSVAQMSEP GSEPCPAEMAQDPDSDAPDQASPOSTQDQKRLRFQLEKIGYIHKDCKIRW ESAYMVQGTSTKVYKFCPRVCRESYNPYRVEDITQSCKRTRCACPVRLRHDPKR PHRQDLCGRKDKRLSCDSTFSFKVIL"
exon	2574..2666
	/gene="Zarl1"
	/number=2
exon	2745..2819
	/gene="Zarl1"
	/number=3
exon	3697..3987
	/gene="Zarl1"
	/number=4
ORIGIN	
Query Match	62.6%; Score 799.4; DB 10; Length 3987;
Best Local Similarity	99.9%; Pred. No. 1.2e-155;
Matches	800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	3 GCGGGCGAGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 62
Db	3 GCGGGCGAGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 62
Qy	63 TCCTTATCGCAGCGCCACCAAGCCGGGATGCTCGAGGTTTCGAGCCAGGGGTGCG 122
Db	63 TCCTTATCGCAGCGCCACCAAGCCGGGATGCTCGAGGTTTCGAGCCAGGGGTGCG 122
Qy	123 ACCGGGCCCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTAGTCGA 182
Db	123 ACCGGGCCCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTAGTCGA 182
Qy	183 CAGCCACACGCGGCGACAGCTCATGGCCCTGTGTGCGGGATGGGTCCCGGTTCGAG 242
Db	183 CAGCCACACGCGGCGACAGCTCATGGCCCTGTGTGCGGGATGGGTCCCGGTTCGAG 242
Qy	243 CAGCCGTGAGCTGCGGTGACGTGAACCGGCGCCGACGCTCGGTGACGTGTTCACT 302
Db	243 CAGCCGTGAGCTGCGGTGACGTGAACCGGCGCCGACGCTCGGTGACGTGTTCACT 302
Qy	303 CGGCGCGCGACCGCTGACGCTCGAGGTTGCCGAGCCAGCCCGGATCGGGTTC 362

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 147464)

# TITLE JOURNAL

## REFERENCE AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,  
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
Collamore, A., Cook, A., Cooke, P., Corum, B., DeKrellano, K.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, B., Erickson, J., Faro, S.,  
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,  
Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,  
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
Meldrim, J., Menees, J., Mihova, T., Miengo, V., Murphy, T., Naylor, J.,  
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (02-FEB-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 2, 2003 this sequence version replaced gi:21206317.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center

# TITLE JOURNAL

## COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: <http://www.seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L17506

Center clone name: 506\_B\_15

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 140214 bases at least Q40  
Consensus quality: 143454 bases at least Q30  
Consensus quality: 144691 bases at least Q20  
Insert size: 144000; agarose-fp  
Insert size: 145464; sum-of-contigs  
Quality coverage: 4.6 in Q20 bases; agarose-fp  
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 653: contig of 653 bp in length  
\* 654 753: gap of 100 bp  
\* 754 1934: contig of 1181 bp in length  
\* 1935 2034: gap of 100 bp  
\* 2035 31996: contig of 29962 bp in length  
\* 31997 32096: gap of 100 bp  
\* 32097 32986: contig of 890 bp in length  
\* 32987 33087: gap of 100 bp  
\* 33087 33811: contig of 725 bp in length  
\* 33812 33911: gap of 100 bp  
\* 33912 35429: contig of 1518 bp in length  
\* 35430 35529: gap of 100 bp

35530 38157: contig of 2628 bp in length  
\* 38158 38257: gap of 100 bp  
\* 38258 41538: contig of 3281 bp in length  
\* 41539 41638: gap of 100 bp  
\* 41639 43241: contig of 1603 bp in length  
\* 43242 43341: gap of 100 bp  
\* 43342 45930: contig of 3589 bp in length  
\* 45931 47030: gap of 100 bp  
\* 47031 51340: contig of 4310 bp in length  
\* 51341 51440: gap of 100 bp  
\* 51441 56135: contig of 4695 bp in length  
\* 56136 56235: gap of 100 bp  
\* 56236 61691: contig of 5456 bp in length  
\* 61692 61791: gap of 100 bp  
\* 61792 70053: contig of 8262 bp in length  
\* 70054 70153: gap of 100 bp  
\* 70154 76663: contig of 6510 bp in length  
\* 76664 86117: contig of 9354 bp in length  
\* 86118 86217: gap of 100 bp  
\* 86219 94174: contig of 7957 bp in length  
\* 94175 94274: gap of 100 bp  
\* 94275 108523: contig of 14249 bp in length  
\* 108524 108623: gap of 100 bp  
\* 108624 123796: contig of 15173 bp in length  
\* 123797 123896: gap of 100 bp  
\* 123897 146961: contig of 23065 bp in length  
\* 146962 147061: gap of 100 bp  
\* 147062 147464: contig of 403 bp in length.  
FEATURES  
Location/Qualifiers  
source  
1..147464  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP24-506B15"  
/clone\_lib="RPC1-24 Male Mouse BAC"  
1..653  
/note="assembly\_fragment"  
misc\_feature  
754..1934  
/note="assembly\_fragment"  
misc\_feature  
2035..31996  
/note="assembly\_fragment"  
misc\_feature  
32097..32986  
/note="assembly\_fragment"  
misc\_feature  
33087..33811  
/note="assembly\_fragment"  
misc\_feature  
33912..35429  
/note="assembly\_fragment"  
misc\_feature  
35530..38157  
/note="assembly\_fragment"  
misc\_feature  
38258..41538  
/note="assembly\_fragment"  
misc\_feature  
41639..43241  
/note="assembly\_fragment"  
misc\_feature  
43342..46930  
/note="assembly\_fragment"  
misc\_feature  
47031..51340  
/note="assembly\_fragment"  
misc\_feature  
51441..56135  
/note="assembly\_fragment"  
misc\_feature  
56236..61691  
/note="assembly\_fragment"  
misc\_feature  
61792..70053  
/note="assembly\_fragment"  
misc\_feature  
70154..76663  
/note="assembly\_fragment"  
misc\_feature  
76764..86117  
/note="assembly\_fragment"  
misc\_feature  
86218..94174  
/note="assembly\_fragment"  
misc\_feature  
94275..108523  
/note="assembly\_fragment"  
misc\_feature  
108624..123796  
/note="assembly\_fragment"







Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokemele, O., Okwuonu, G., Olatunmbi, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Primus, E., Pu, L.-L.,  
Plopper, F., Polindexter, A., Popovic, D., Reeves, K., Regier, M.A., Reigh, R.,  
Ruiz, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,  
Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Williams, R., Wlezyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 235390)  
Worley, K.C.  
Direct Submission  
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 235390)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 20, 2002 this sequence version replaced gi:23195602.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GZCZ  
Center clone name: CH230-159N5  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 216235 bases at least Q40  
Consensus quality: 218057 bases at least Q30  
Consensus quality: 219408 bases at least Q20  
Estimated insert size: 221324; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
\*  
source  
1 235390: contig of 235390 bp in length.  
Location/Qualifiers  
1. 235390  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-159N5"  
1. 1856  
/notes="wgs\_end\_extension  
clone\_end:T7"  
complement(4755..38027)  
/notes="clone\_boundary  
clone\_end:T7  
site:  
end\_sequence:BH321013"  
complement(227221..228048)  
/notes="clone\_boundary  
clone\_end:Sp6  
site:  
end\_sequence:BH321015"  
232121..235390  
/notes="wgs\_end\_extension  
clone\_end:Sp6"  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
ORIGIN  
Query Match 49.2%; Score 628; DB 2; Length 235390;  
Best Local Similarity 88.1%; Pred. No. 4.6e-120; Indels 6; Gaps 2;  
Matches 708; Conservative 0; Mismatches 90;  
QY 3 GGCGGCGGAGCGCGGAGCGCACCCATGTTCCCGGGGAGCAGCCTTCACCCCTGCCGCGCA 62  
DB 27469 GGCGGCGGAGCGCGGAGCGCACCCATGTTCCCGGGGAGCAGCAGCCACCCATGCGCGCA 27410  
QY 63 TCCTTATTCGCG-----AGGCCACCAAGCCGGGAGTGTGGAGGTTCCGAGCCAGGGGCTG 119  
DB 27409 TCCTTATTCGCGCGCGGAGCGCACCCATGTTCCCGGGGAGGTTGGAGGTTTGGAGCAGGGGCTG 27350  
QY 120 CGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179  
DB 27349 CAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27290  
QY 180 CGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239  
DB 27289 TGACAGCTATCAGCGAGCGCAGCTCATGCGCTTGTCTGTGGAGATGGTTCCTCCGCGCGGT 27230  
QY 240 CAGCAGCGGTGAGCGCTGCGGTTGAGTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299  
DB 27229 CAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27170  
QY 300 ACTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359  
DB 27169 GCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27110  
QY 360 TTCTCTGTCAACCCCGGTGGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419  
DB 27109 TTCTCTGTCAACCCCGGTGGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27050  
QY 420 AGCCCGCGTTCCTGCGTGCAGCTTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479  
DB 27049 CGCCCTGTACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26990  
QY 480 CAGGCGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539  
DB 26989 CAGGCGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26930  
QY 540 GCGCGAGAGGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599  
DB 26929 GCGCGAGAGGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26870  
QY 600 TCAGGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659  
DB 26869 CCAGGCTGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26813



```

Qy 660 GCGGCGGATGAGTCTGAGCTGGAGCGAGGAGCCATGTCTGTCGCGAGAGATGGCTCA 719
    |||||||
Db 2681.2 GCGGCGGATGAGTCTGAGCTGGAGCGAGGAGCCATGTCTGTCGAGATGGCTCA 26753
    |||||||

Qy 720 GGACCCCGGTGATTCGGATGCCCTCGAGACGAGCCCTCCCGGAAAGCAGCGAGCAGGA 779
    |||||||
Db 26752 GGACCCCGGTGATTCGGATGCCCTCGAGACGAGCCCTCCCGGAAAGCAGCGAGCAGGA 26693
    |||||||

Qy 780 CAAGGAGCGCTGGCTTTCCAGTT 803
    |||||||
Db 26692 CAAGGAGCGCTGGCTTTCCAGTT 26669
    |||||||

RESULT 8
AC125993/c
LOCUS Rattus norvegicus clone CH230-74L11, *** SEQUENCING IN PROGRESS ***
DEFINITION AC125993
VERSION AC125993.3 GI:30522839
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 237695)
Muzny,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Dengson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorjse,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuheva,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Naunkervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackeleme,O., Okwuonu,G., Olarpunsgoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanckoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,S., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,D., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,S., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

```

```

TITLE JOURNAL
REFERENCE JOURNAL
AUTHORS JOURNAL
TITLE JOURNAL
REFERENCE JOURNAL
AUTHORS JOURNAL
TITLE JOURNAL
COMMENT

```

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 237695)

Worley, K. C.

Direct Submission

Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 237695)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center of Medicine

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GGIZ

Center clone name: CH230-74L11

----- Summary Statistics

Assembly program: Atlas 3.0:

Consensus quality: 229641 bases at least Q40

Consensus quality: 231444 bases at least Q30

Consensus quality: 232679 bases at least Q20

Estimated insert size: 242796; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced

\* By the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 237695: contig of 237695 bp in length.

```

FEATURES
    source
        1..237695
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-74L11"
    misc_feature
        1..1905
            /note="wgs end extension
            clone end:T7"
    misc_feature
        complement(5394..6172)
            /note="clone boundary
            clone end:T7
            site:ECORI
            end sequence:BH339391"
    misc_feature
        complement(234275..235086)

```



```
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end sequence:BH339393"
```

## ORIGIN

[illegible]

## RESULT 9

AC127083/C  
LOCUS AC127083 262139 bp DNA linear  
DEFINITION Rattus norvegicus clone CH230-69f8, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
4 unordered pieces.  
ACCESSION AC127083  
HTG 20-NOV-2002

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE

### AUTHORS

AC127083.3 GI:24635317  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)

1 (bases 1 to 262139)

AC127083.3 GI:24635317  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 262139)  
 Muzny,D,Marie, Metzker,M, Lee,, Abramzon,S,, Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
 Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,  
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,  
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,J.,  
 Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lorenshewa,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J.,  
 Maheshwari,M., Mahindaratne,M., Mahmood,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Mesnen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,  
 Nwaekemele,O., Okwuonu,G., Olarunpsagoon,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,  
 Plopper,F., Poindexter,A., Popovic,D., Prims,E., Pu-L-L.,  
 Puazo,M., Quintero,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
 Sanders,W., Savary,G., Scheraz,S., Scott,G., Shatsman,S., Shen,H.,  
 Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,  
 Shedd,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svingey,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstein,G. and Gibbs,R.A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

**Wellstock, G.: and  
Direct Submission**

2 (bases 1 to 262139)

Worley, K.C.

## Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 262139)

Rat Genome Sequencing Consortium.

**TITLE**

**JOURNAL**

**JOURNAL**  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads



Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 212848)

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Ton,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Mihova,T., Mlenea,V., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 28, 2002 this sequence version replaced gi:18450109.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L21014

Center clone name: 300\_K\_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200069 bases at least Q40

Consensus quality: 206302 bases at least Q30

Consensus quality: 208218 bases at least Q20

Insert size: 209348; sum-of-contigs

Quality coverage: 5.7 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1

\* 641: contig of 841 bp in length

\* 741: gap of 100 bp

\* 1427: contig of 686 bp in length

\* 1428: gap of 100 bp

\* 1527: contig of 1643 bp in length

\* 3170: contig of 100 bp

\* 3171: gap of 100 bp

\* 3270: contig of 1126 bp in length

\* 4396: contig of 100 bp

\* 4496: gap of 100 bp

\* 5909: contig of 1413 bp in length

\* 5910: gap of 100 bp

\* 7583: contig of 1574 bp in length

\* 7684: gap of 100 bp

\* 8975: contig of 1292 bp in length

\* 9075: gap of 100 bp

\* 9076: contig of 1182 bp in length

\* 10257: gap of 100 bp

\* 10358: contig of 1251 bp in length

\* 11608: gap of 100 bp

\* 11709: contig of 1581 bp in length

\* 13289: gap of 100 bp

\* 13390: contig of 1946 bp in length

\* 15335: gap of 100 bp

\* 15435: contig of 815 bp in length

\* 16250: gap of 100 bp

\* 16351: contig of 1319 bp in length

\* 17689: gap of 100 bp

\* 17770: contig of 2138 bp in length

\* 19907: gap of 100 bp

\* 20007: contig of 1463 bp in length

\* 21470: gap of 100 bp

\* 21570: contig of 1516 bp in length

\* 23086: gap of 100 bp

\* 23186: contig of 1537 bp in length

\* 24723: gap of 100 bp

\* 24823: contig of 2923 bp in length

\* 27746: gap of 100 bp

\* 27847: contig of 2027 bp in length

\* 29873: gap of 100 bp

\* 33212: contig of 3239 bp in length

\* 33312: gap of 100 bp

\* 33313: contig of 28420 bp in length

\* 61732: gap of 100 bp

\* 61733: contig of 1960 bp in length

\* 63732: gap of 100 bp

\* 63793: contig of 3339 bp in length

\* 67231: gap of 100 bp

\* 67331: contig of 3591 bp in length

\* 70922: gap of 100 bp

\* 71022: contig of 5908 bp in length

\* 76930: gap of 100 bp

\* 77030: contig of 5593 bp in length

\* 82623: gap of 100 bp

\* 82723: contig of 5152 bp in length

\* 87875: gap of 100 bp

\* 87975: contig of 5799 bp in length

\* 93774: gap of 100 bp

\* 93775: contig of 10331 bp in length

\* 93875: gap of 100 bp

\* 104205: contig of 100 bp

\* 104306: gap of 100 bp

\* 119104: contig of 14799 bp in length

\* 119204: gap of 100 bp

\* 132299: contig of 13095 bp in length

\* 132300: gap of 100 bp

\* 13400: contig of 16571 bp in length

\* 148970: gap of 100 bp

\* 149071: contig of 16644 bp in length

\* 165715: gap of 100 bp

\* 165815: contig of 17555 bp in length

\* 183369: gap of 100 bp

\* 183469: contig of 26822 bp in length

\* 210291: gap of 100 bp

\* 210292: contig of 2457 bp in length.

\* 210392: Location/Qualifiers

FEATURES



Qy		414	GACCGTAGCCCGGCTTCGTCCGTGGTGA	CCTTCTGTGGCGCTCTCTCCTCACTGGAGGTTCG	473
Db		531	CACCGTGCGCGTGACTGCGCCCTGGCCT	TTCGGCGCTCTCACCGCTTCTGTGGAGGGGCC	590
Qy		474	GGGAGGC-----	AGCGACACCCACGAAGGAGGGGAGCCCGGCATC	518
Db		591	CGGCGCCCGCGGGCGGAGCAGAGGT	CCGGGGCGTTCGACCGAGAGAGGGGGCGCCGCGCC	650
Qy		519	CTCGGGGACCCGGGAAACCGGAGCCGAGAG	GTGGCCCGCAGGAAAGCGGTCCCCCAGACC	578
Db		651	CGCGCGGCTTAACAGCCACAGAGGAGGG	AGGTGTGGACGAAGAAGCGCCCCCGCGCGCC	710
Qy		579	GC----	GAAACGAGAGAGGGCGATGTTTCAGGCT	628
Db		711	GCAGTCGCGACACGCGAGCGAGGGCCC	AGGCCGCGAGTCCGAGCGAGCTGGGAGCAGCGCGC	770
Qy		629	-----	AGCAGACCCACCGGAGGACCGGAAACAGT	677
Db		771	CGACGTTCCGAGTGTCCGCCCGCAGAGG	CCACAGGCGCGAGCGCGTCCGCGGTCCGC	830
Qy		678	GCCTGGGAGCGAGGAGCCATTGCTCTG	CCGACAGATGGCTCAGGACCCCGGTGATT----	733
Db		831	GCTAAGGAGCCCGGGCAAACCTCCGT	CGSGGGAGGGCCCGACGCGCGCGACGAGCG	890
Qy		734	-----	CGAGTCCCTCGAGACACAGGCCCTCCCG	788
Db		891	GGAGGCGGCGGTTCGCGGAGAGGGGCG	CGGTTCGCCACGGAGCTGGCAGAGGAGCG	950
Qy		789	CTTSCGTTTTCCAGTTCTTAGAGCAGA	GTACGGCTACTATCATCTGCAAGGACTGCAAAAT	848
Db		951	GCTTGCCTTCCAGTTCTTAGAGCAGAA	ATATGGCTATTACACATGCAAGGACTGCAACAT	1010
Qy		849	CCGCTGGGAGAGCGCCTATGTGTGGT	GTGTGCAGGGCACCACTAAGGTGTACTTTCAAAC	908
Db		1011	CCGCTGGGAGAGTCTTATGTGTGTGT	GTGTACAGGGAACCTACAAGGT-TTACTTTCAAAC	1069
Qy		909	AGTTCTGCGAGTGTGTAGAAAATCC	TACAAACCTTACAGAGTGGAGGACATCACCTGTC	968
Db		1070	AGTTTTGCAGAACTTGTGAGAACTT	TATAACCCCTTACCGAGTGGAGGATATCACCTGTC	1129
Qy		969	AAGTTTGTAAAAAGAACTTAGATGT	GCCTTGCCTCAGTTCAGATTTCGCCACGTGGACCTTAAAC	1028
Db		1130	AAAGTTGTAAACAAACAGAGATGTT	CTCTGSCCAGTAAAAAATTCGCACAGTGGACCTTAAAC	1189
Qy		1029	GCSCCCATCGGCAAGACTTGTGTGG	GAGATGCAAGGACAAACGCTCTCTGGGACAGCA	1088
Db		1190	GGCCCCACCGTCAAGATTTGTGCG	GTAGATGCAAAAGCAAACGCTCTCTGTGACAGCA	1249
Qy		1089	CTTTCAGCTTCAAATACATCATTTAG		1114
Db		1250	CTTTTCAGCTTCAAATACATCATTTAG		1275

RESULT 12	AC108848/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
			AC108848	212848 bp	DNA	linear	HTG 28-APR-2002
			Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered pieces.				
			AC108848				
			AC108848.2	GI:203336129			
			HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.				
			Mus musculus (house mouse)				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 212848)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP23-300K5  
Unpublished  
2 (bases 1 to 212848)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,E.,

Choepl,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,  
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,  
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,  
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
Retka,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 212848)

Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campolano,A., Chang,J.,  
Chazaro,B., Choepl,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,  
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,  
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,  
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retka,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 28, 2002 this sequence version replaced gi:19450109.

All repeats were identified using RepeatMasker:  
Smith,A.F.A. & Green,P. (1996-1997)

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Role of the Parent in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Impact of the School Environment on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the Student in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of the Teacher's Personality on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Role of the School in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Impact of the Community on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the Society in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of the Culture on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100



893 CTTCAAGTATATCAATT 909

RESULT 14  
 AY283177  
 LOCUS AY283177 963 bp mRNA linear VRT 21-AUG-2003  
 DEFINITION Takifugu rubripes zygote arrest 1 (Zarl) mRNA, complete cds.  
 ACCESSION AY283177  
 VERSION AY283177.1 GI:30908936  
 KEYWORDS  
 SOURCE Takifugu rubripes (Fugu rubripes)  
 ORGANISM Takifugu rubripes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontoidea; Tetraodontidae; Takifugu.  
 1 (bases 1 to 963)  
 Wu, X., Wang, P., Brown, C.A., Zilinski, C.A. and Matzuk, M.M.  
 Zygote arrest 1 (Zarl) is an evolutionarily conserved gene  
 expressed in vertebrate ovaries  
 Biol. Reprod. 69 (3), 861-867 (2003)  
 22811438  
 12773403  
 2 (bases 1 to 963)  
 Wu, X. and Matzuk, M.M.  
 Direct Submission  
 Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 Location/Qualifiers  
 1..963  
 /organism="Takifugu rubripes"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:31033"  
 1..963  
 /gene="Zarl"  
 1..963  
 /gene="Zarl"  
 /note="maternal factor"  
 /codon\_start=1  
 /product="zygote arrest 1"  
 /protein\_id="AAP37039.1"  
 /db\_xref="GI:30908937"  
 /translation="MATYCDPEVDSYFYSNPMGRYPHRDAGWKYKSYLSHYGDT  
 SFASNQRAQLKLSIQNPLRLKANTKDVAQVNPKRDAQVQCSIGRPTLLV  
 VKRELRLKPLNGPGTQKTEGEVRYPTILAVSPIAFRSVTSFLVETGKDRPAE  
 AQAEELPQGPQKGGENQAGETNANLPQRKQSEDAQTADAGSGKARVRQF  
 LEQKYGYHCRNLFWESAYWCVQGNKYFKQFCRCKQKDFNPRVEDITCHVCN  
 KARCACAEYQRHVDPRPRHQDLGRCKRGLSCDSTFSFKYIV"  
 ORIGIN  
 Query Match 15.3%; Score 194.8; DB 5; Length 963;  
 Best Local Similarity 75.4%; Pred. No. 4.7e-30;  
 Matches 255; Conservative 0; Mismatches 82; Indels 1; Gaps 1;  
 QY 776 AGCAAGAGCGCCGTTCAGTTCCTTAGAGCAGAAGTACGGCTACTACTGCA 835  
 Db 626 AGGCAAGCGCGTGTCCGCTCCAGTTCTCGAACAGAAAGTAGGGCTACTACTGCA 685  
 QY 836 AGGACTGCAAAATCCGGTGGAGAGCGCTATGTGTGGTGTGTGCAGGGCACCAGTAAGG 895  
 Db 686 GAGAATGCAACTTGCATGGGAGAGCGCGTACGTTTGGTGGCTTCAGGGCACTAACAGG 745  
 QY 896 TGTACTTCAAAACAGTTCTGCCGAGTGTGTGAAATCTTACACCCCTTACAGAGTGGAG 955  
 Db 746 T-TTACTCAAGCAGTTCTGTAGGAATGCCAAAAGACTTTAAACCGTAGCCGTAGAG 804  
 QY 956 GACATCACTGTCAAAGTTGTAAAAAGAACTAGATGTGCTGCCCACTCAGATTTCGCCAC 1015  
 Db 805 GACATCACATGTCACTATGCAACAGGCGCGTGTGCTCGCGAGAAACGCAAGGCCAC 864  
 QY 1016 GTGACCTTAACGCCCCCATCGGCAAGACTTGTGTGGAGATGCAAGCAAAACGCTTG 1075  
 Db 865 GTTGACCCAAAGAGGCCCAAGGAGGACTTGTGCGGAGTGAAGGCAAGCGGCTG 924

ORIGIN	Query Match	15.2%	Score 194.6;	DB 5;	Length 1084;
	Best Local Similarity	75.1%;	Pred. No. 5.2e-30;		
	Matches 256;	Conservative 0;	Mismatches 84;	Indels 1;	Gaps 1;
QY	774	GCAGGCAACAGGAGCGCTCGCTTTCCAGTTCTTTAGACGAGAAGTACCGGTACTATCACTG	833		
DB	692	GAAGTCCAAGGCTCGTGTGAGATTTCAGTCTTTGGAGCAGAAGTATGGATTTCTATCATTTG	751		
QY	834	CAAGGACTGC AAAATCCGGTGGGAGAGGCGCTATGTGTGTGTGTGTGTCAGGGCACCAGTAA	893		
DB	752	CAAAAGACTGC AACCTACGGTGGGAAAGTGCCTTATGTGTGTGTGTGTCCAGGAACAACAA	811		
QY	894	GGTGTACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCCTACAAACCCCTTTACAGAGTGG	953		
DB	812	GST-TTATTTCAAGCAGTTCTGCAGAACATGCCAGAAATCATTTCAACCCATACCGGGTTG	870		
QY	954	AGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAAGATTTTCGCC	1013		
DB	871	AGGACATAGCATGTTCAGACTTTGCAGAAAGCTCCGTCGACATGTTCTGTCAAGTCGCGTC	930		
QY	1014	ACGTGGACCTTAAACCGCCCCCATCGGCAAGACTTTGTGTGGAGATGCAAGGACAAACGCC	1073		



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 13:40:31 ; Search time 577 Seconds  
(without alignments)  
9401.991 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagcggcgagcgcgaggga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002s:\*  
7: Geneseq2003as:\*  
8: Geneseq2003bs:\*  
9: Geneseq2003cs:\*  
10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1265	99.1	1276	3	Aad00294 Mouse ooc
2	1234.2	96.6	1258	7	Abz24589 Mouse ova
3	799.4	62.6	6873	7	Abz24590 Mouse ova
4	753.8	59.0	4090	7	Abz24591 Mouse ova
5	134.6	10.5	2075	7	Abz24592 Human ova
6	58.8	4.6	114955	2	Aax53491 Human ade
7	57.8	4.5	2000	7	Ada71938 Rice gene
8	54.2	4.2	1032	7	Ada69587 Rice gene
9	52	4.1	114955	2	Aax53491 Human ade
10	51.6	4.0	987	7	Abz37562 Streptomy
11	51.6	4.0	59816	7	Abz37516 Streptomy
12	51.6	4.0	59816	7	Abz37515 Streptomy
13	49.6	3.9	400	7	Aav44436 Mycobacte
14	49.6	3.9	400	2	Aav64545 M. tuberc
15	49.6	3.9	400	2	Aaz19134 M. tuberc
16	49.6	3.9	400	2	Aaz19134 M. tuberc
17	49.6	3.9	14800	6	Ab166291 Lung canc
18	49.6	3.9	14835	6	Aas94858 Human DNA
19	49	3.8	7175	2	Aa94658 Human neu
20	49	3.8	7175	2	Aa94658 Human neu
21	49	3.8	7175	3	Aaa71704 DNA encod
22	49	3.8	7175	3	Aaa71704 Human cal
23	49	3.8	7177	7	Abz58367 Human N-t

24	49	3.8	7266	2	AAV29059 Human cal
25	49	3.8	7362	2	AAQ37817 Sequence
26	49	3.8	7362	2	AAQ84657 Human neu
27	49	3.8	7362	2	AAV42685 DNA encod
28	49	3.8	7362	3	AAV71703 Human cal
29	49	3.8	7362	3	AAQ39955 Human cal
30	49	3.8	7364	6	ABL65869 Lung canc
31	49	3.8	7364	7	ABZ58366 Human N-t
32	49	3.8	7376	2	AXX88001 N-type ca
33	49	3.8	109519	5	AAS08693 Micromono
34	48.2	3.8	524	4	AAK92108 Human CDN
35	48.2	3.8	524	4	AAK93547 Human CDN
36	48.2	3.8	1769	4	AAK94827 Human ful
37	48	3.8	110000	4	Continuation (40 o
38	47.6	3.7	5944	6	ABL32253 Human imm
39	47.6	3.7	5944	6	ABL92203 Chemical
40	46.8	3.7	536	9	ADB68842 Minority
41	46.8	3.7	1383	7	ADA71056 Rice gene
42	46.6	3.6	745	2	AAQ62837 GGF gene
43	46.6	3.6	745	2	AAQ58330 GGF gene
44	46.6	3.6	745	2	AAQ30988 Human gli
45	46.6	3.6	745	2	AAQ48077 Human gli

## ALIGNMENTS

### RESULT 1

AD000294  
ID AD000294 standard; cDNA; 1276 BP.  
XX  
AC AD000294;  
XX  
DT 05-SEP-2000 (first entry)  
XX  
DE Mouse oocyte-specific O1-180 cDNA clone.  
XX  
KW Oocyte-specific; ovary; O1-180; mouse; gynaecological; treatment; screen;  
KW cell proliferative disorder; cell degenerative disorder; contraceptive;  
KW modulator; signalling pathway; human infertility; cancer; ovulation; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 28..1113  
FT /\*tag= a  
FT /product= "Mouse oocyte-specific protein, O1-180"  
XX  
PN WO200024755-A1.  
XX  
PD 04-MAY-2000.  
XX  
PF 28-OCT-1999; 99WO-US025209.  
XX  
PR 28-OCT-1998; 98US-0106020P.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Matzuk MM, Wang P;  
XX  
DR WPI; 2000-350684/30.  
XX  
DR P-PSDB; AAY70948.  
XX  
PT O1-180, O1-184 and O1-236 polypeptides and nucleic acids encoding them,  
XX useful for evaluating potential contraceptives to block ovulation in a  
XX reversible manner.  
XX  
PS Claim 2; Fig 1; 54pp; English.  
XX  
CC The present sequence is the cDNA encoding the mouse oocyte-specific  
XX protein O1-180, expressed in the oocytes of primary (one-layer) prenatral  
XX follicles through ovulation. It provides in vitro and in vivo reagents  
XX for studying ovarian development and function. This sequence has









Claim 2; Page 136-137; 141pp; English.

The present sequence is that of a human ovary-specific O1-180 polynucleotide. In mice, loss of O1-180 is associated with female infertility and subfertility. The invention provides ovary-specific and oocyte-specific murine and human O1-180, O1-184 and O1-236 polynucleotides and polypeptides. These genes and their protein products appear to relate to various cell proliferative or degenerative disorders, especially those involving ovarian tumours, such as germ line tumours and granulosa cell tumours, or infertility, such as premature ovarian failure. The invention provides a method for detection of a cell proliferative or degenerative disorder of the ovary, which is associated with the expression of O1-180, O1-184 or O1-236. It also provides a method for treating such disorders by using an agent which suppresses or enhances the respective activities of O1-180, O1-184 or O1-236, and a method of screening for compounds that interact and/or modulate the expression or activity of the ovary-specific genes. These compounds are possible contraceptive or fertility enhancing agents. The modulator is preferably a polypeptide, small molecule or polynucleotide sequence

Sequence 2075 BP; 588 A; 413 C; 473 G; 601 T; 0 U; 0 Other;

	Query Match	10.5%;	Score 134.6;	DB 7;	Length 2075;
	Best Local Similarity	71.1%;	Pred. No. 6.6e-22;		
	Matches 207;	Conservative 0;	Mismatches 79;	Indels 5;	Gaps 2;
Qy	968	CAAAAGTTGTAAAGAAGCTAGATGTCCTGCCAGTTCAGATTGCCACGTGGACCTTAAA	1027		
Db	1360	CAGAGTTGTAAACAAACGAGATGTTCTGTGCCACGTAAACCTTGCCACGTGGACCTTAAA	1419		
Qy	1028	CGCCCCCATCGGCAAGACTTGTGTGGGAGATGCAAGCAACAACGCTGTCTCTGGACAGC	1087		
Db	1420	CGGCCCCACCGTCAAGATTGTGTGGGTAGATGCAAGCAACAACGCCGTCTCTGTGACAGC	1479		
Qy	1308	ACCTTTCAGCTTCAAATACATCATTTTA-GTGAGAGTCGAAAACGTTTCTGTGTAGTGGGGC	1146		
Db	1480	ACTTTCAGCTTCAAATACATCATTTTAGTGTGAAGTCAGTGTGCTGTGCATGCCCTGATG	1539		
Qy	1147	TAATGGAATGCAACAGTGAGCTTTCTCCCTCTTTCACCTCTTCCTCTTCCAAATCTTCA	1206		
Db	1540	GAGTAGACGAGTGAGCTTTTCCGTGCCCTTCTCCTCACCTTCCCTTCTCAAAATACTTCA	1599		
Qy	1207	TGACAGACAGGTGTTACTTGGGATATAAAGCCCTGTGAATAAAAAGGTATTGCAA	1257		
Db	1600	TGAAGGCGAGTGTATTCTG-----AAAGAGCTTCAATAAAGGTATTGCAA	1646		

## RESULT 6

AA53491  
ID AAX53491 standard; DNA: 114955 BP.

AA  
AC AAX53491:

AA	05-JUL-1999	(first entry)
DT		

XX Human adenosine A1 receptor antisense oligonucleotide fragment.

Antisense oligonucleotide; multiple target; antisense treatment;  
 impaired respiration; inflammation; lung disease;  
 pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 acute asthma; allergy; asthma; impeded respiration;  
 respiratory distress syndrome; pain; cystic fibrosis;  
 pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
 colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 prostate cancer; ss.

AA OS Synthetic.

AA WO9913886-A1.  
PN

25-MAR-1999.

Qy 421 GCCCGTTCTCGTCCGTGACCTTCTGTGGCCTCTCCTCCTCACTGGAGGTTCGGGAGGC 480

XX	17-SEP-1998;	98WO-US019419.
XX		
XX	17-SEP-1997;	97US-0059160P.
XX	PR	
XX	09-JUN-1998;	98US-00093972.
XX		
XX	(UYEC-) UNIV EAST CAROLINA.	
XX		
XX	Nyce JW;	
XX		
XX	WPI; 1999-229400/19.	
XX		
XX	New antisense oligonucleotides used in treatment of, e.g. pulmonary	
XX	vasoconstriction.	
XX		
XX	Disclosure; Page 37; 120pp; English.	
XX		
XX	The specification describes antisense oligonucleotides (AAx52869-X55271)	
XX	directed against at least 2 mRNAs selected from target genes, coding and	
XX	non-coding regions of RNAs corresponding to target genes, gene initiation	
XX	codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'	
XX	-end and the juxta-section between coding and non-coding regions and all	
XX	segments of RNAs encoding proteins associated with one or more diseases,	
XX	conditions or mixtures. The antisense oligonucleotides may be derived	
XX	from sequences AAx5272-74. These multiple target oligonucleotides	
XX	(specifically AAx55180-271) can be used for the antisense treatment of	
XX	diseases and conditions. Typical diseases and conditions are those	
XX	associated with impaired respiration and inflammation, including lung	
XX	diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,	
XX	acute asthma, allergies, asthma, impeded respiration, respiratory	
XX	distress syndrome, pain, cystic fibrosis, pulmonary hypertension,	
XX	pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary	
XX	disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.	
XX	colon cancer, breast cancer, lung cancer, pancreatic cancer,	
XX	hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as	
XX	well as all types of cancers which may metastasize or have metastasized	
XX	to the lungs, including breast and prostate cancer	

Sequence 114955 BP: 6071 A: 29417 C: 36712 G: 21328 T: 0 U: 21427 Other: 0

Query Match	4.6%	Score 58.8	DB 2	Length 114955
Best Local Similarity	32.6%	Pred. No. 0.0025		
Matches 251: Conservative	71	Mismatches 441	Indels	8 Gaps



PN W02003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX  
XX PF (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
XX WPI; 2003-175290/17.  
XX  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX  
XX Claim 6; SEQ ID NO 2910; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
XX  
XX Sequence 1032 BP; 189 A; 331 C; 394 G; 118 T; 0 U; 0 Other;  
SQ

Query Match 4.2%; Score 54.2; DB 7; Length 1032;  
Best Local Similarity 43.4%; Pred. No. 0.007;  
Matches 251; Conservative 0; Mismatches 328; Indels 0; Gaps 0;

QY 137 CTTTCTTCCCGGTACAGACAGTCTATGCGCGGAGTACTGTCAGACCCACGCGG 196  
Db 167 CTTTCTTCCCGGTACAGACAGTCTATGCGCGGAGTACTGTCAGACCCACGCGG 226  
QY 197 CACAGCTATGCGTCTGTCGCGGAGTGGTCCCGGTGCTGAGCGCGTGCAGCTG 256  
Db 227 AGAAGAGCGCGCTTACCGCGGAGCAGTGCATCTGCTGAGAGGAGCTTCGAGGAG 286  
QY 257 CGGTGCGAGTGAAACCGCGCGCGAGCGCTCGGTGCGAGTGTTCATCTCGGCGCGCACGC 316  
Db 287 AGAACAAGCTGAGCGCGGAGCGAAGACGAGCTGGCGCGGAAGCTAGGGCTGCAGCCGC 346  
QY 317 TCGAGCTCGAGGTGCCAGCGCGCGAGCGCGCGAGTTCCTGTCACACCCCGTG 376  
Db 347 GCGAGTTCGCGGTGTTCCAGAACCGCGCGCGGTGGAAGACCAAGCAGCTCGAGC 406  
QY 377 GCGAGCGCGCGCGGAGTCCCGCGATCTCGGCGAGACCGTAGCCCGTTCGTGTCG 436  
Db 407 GCGACTTCAGCGCTCAGGAGTGTTCGAGCGCTTCGCGCGCGAGCGCGCTCC 466  
QY 437 TGACCTTCTGTGCGCTTCTCTCTACTGAGGTTGCGGAGCGAGCGAGCACCCACGA 496  
Db 467 TCCAGGACAAACACCGCTCCACTCTCAGGTCTATGCTGTTGACCGAGAGCTGCAAGA 526  
QY 497 AGGAGAGGGAGCGCGGATCTCTGCGGACCGGAGCGCGAGAGGTGCGCG 556  
Db 527 AGAGACGACGACCGAGGAGCGAGCGCGCGCGCGCTTTCAGCTGCTGCTGCGG 586  
QY 557 CGAGGAAGCGGTCCCGCGCGCGAGGAGGAGCGAGTGTTCAGGCTGCGAGGCGAG 616  
Db 587 CGGCGGACGTGAGGTGCGCGTCCCGGACCGCGAGGACCGGCGCTGGAGGAGCGCGG 646  
QY 617 CCGGGTGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCTG 676

Db 647 CGGCGTTTCGAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 706  
QY 677 AGCCTGGGAGCGAGGAGCGCATGTCTGCGCGGAGAGATGG 715  
Db 707 GCGCGGAGCGCGGTGTTGGACACGAGCGGCGCACTGG 745

RESULT 9  
AA53491/C  
ID AAX53491 standard; DNA; 114955 BP.  
XX  
XX AAX53491;  
XX  
XX 05-JUL-1999 (first entry)  
XX  
XX Human adenosine A1 receptor antisense oligonucleotide fragment.  
XX  
XX Antisense oligonucleotide; multiple target; antisense treatment;  
XX impaired respiration; inflammation; lung disease;  
XX pulmonary vasoconstriction; inflammation; allergic rhinitis;  
XX acute asthma; allergy; asthma; impeded respiration;  
XX respiratory distress syndrome; pain; cystic fibrosis;  
XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
XX colon cancer; breast cancer; lung cancer; pancreatic cancer;  
XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
XX prostate cancer; ss.  
XX  
XX Synthetic.  
XX  
XX WO913886-A1.  
XX  
XX 25-MAR-1999.  
XX  
XX 17-SEP-1998; 98WO-US019419.  
XX  
XX 17-SEP-1997; 97US-0059160P.  
XX  
XX 09-JUN-1998; 98US-00093972.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX  
XX Nyce JW;  
XX  
XX WPI; 1999-229400/19.  
XX  
XX New antisense oligonucleotides used in treatment of, e.g. pulmonary  
XX vasoconstriction.  
XX  
XX Disclosure; Page 37; 120pp; English.  
XX  
XX The specification describes antisense oligonucleotides (AAX52869-X55271)  
XX directed against at least 2 mRNAs selected from target genes, coding and  
XX non-coding regions of RNAs corresponding to target genes, gene initiation  
XX codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'  
XX -end and the 3'-end of the RNA-section between coding and non-coding regions and all  
XX segments of the RNAs encoding proteins associated with one or more diseases,  
XX conditions or mixtures. The antisense oligonucleotides may be derived  
XX from sequences AAX55180-271. These multiple target oligonucleotides  
XX (specifically AAX55180-271) can be used for the antisense treatment of  
XX diseases and conditions. Typical diseases and conditions are those  
XX associated with impaired respiration and inflammation, including lung  
XX diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,  
XX acute asthma, allergies, asthma, impeded respiration, respiratory  
XX distress syndrome, pain, cystic fibrosis, pulmonary hypertension,  
XX pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary  
XX disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.  
XX colon cancer, breast cancer, lung cancer, pancreatic cancer,  
XX hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as  
XX well as all types of cancers which may metastasize or have metastasized  
XX to the lungs, including breast and prostate cancer  
XX  
XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;  
SQ









Mon Sep 27 09:07:51 2004

```
XX 17-FEB-1999; 99WO-US003265.
XX
XX 18-FEB-1998; 98US-00024753.
XX
XX 05-MAY-1998; 98US-00072596.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
XX WPI; 1999-527416/44.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis.
XX
XX Claim 4; Page 215-216; 323pp; English.
XX
XX This invention describes novel recombinant antigens and their encoding
XX nucleic acids derived from Mycobacterium tuberculosis. The novel
XX polypeptides are useful for detecting M. tuberculosis infection in a
XX biological sample by detecting antibodies which bind with the
XX polypeptides, and are useful as vaccines for immunizing against M.
XX tuberculosis infection. The new detection methods are needed as current
XX vaccination strategies do not provide 100% immunity
XX
XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 U; 0 Other;
XX
Query Match 3.9%; Score 49.6; DB 2; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.064;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 458 CCTCACTGGAGGTTCGGGAGGCGAGCAGACACCCACGAGGGAGGGAGGCCGGCAT 517
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 CTTCAACCAAGCGCGCGCGGCAACCGCGGCAACCGCGGCGGCGGCGGCGGCGGCGG 178
QY 518 CCTCGGGGACCGGGAAACCGAGAGGAGGTGGCGGAGGAAAGCGGTCCCCAGC 577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 CGCGGGAACCGCGGGAACCGCGGAGACACACACACCGCGCGCGCGCGGCGGCGGCGG 238
QY 578 CGCGAAGCGAGGAGGGCGATGTTTCAGGTCAGGGGAGCGCGGCGGCGGCGGCGGCGG 637
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 CGGCGACGGCGGGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGG 298
QY 638 CACCGGAGGACCGGAACAGTGTGGCGCGATGCGAGCTGAGCTGGAGCGAGGAGCCAT 697
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 CGGCACCGCGCGGCGGCAACAGGCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 358
QY 698 GTCTGCGCGAGAGATGGCTCAGG 721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 CGGCGGCGAGGTGCACTCTCAGG 382
```

Search completed: September 24, 2004, 17:34:43  
Job time : 587 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 17:17:47; Search time 114 Seconds  
(without alignments)  
6216.424 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aagggcgggcgggcgggga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	4.2	5467	1	US-07-745-206A-12
2	54	4.2	5467	2	US-08-311-363-12
3	53.2	4.2	7218	1	US-08-232-463-14
4	50.8	4.0	6232	3	US-08-456-200B-11
5	49.6	3.9	400	4	US-09-056-556-179
6	49.6	3.9	400	4	US-09-072-596-174
7	49.6	3.9	400	4	US-09-072-967-179
8	49	3.8	7175	1	US-08-455-543A-8
9	49	3.8	7175	2	US-08-133-078B-8
10	49	3.8	7175	2	US-08-223-305C-8
11	49	3.8	7175	3	US-08-149-097D-8
12	49	3.8	7175	3	US-08-949-386-8
13	49	3.8	7175	3	US-08-450-562-8
14	49	3.8	7175	4	US-08-984-709A-8
15	49	3.8	7175	4	US-08-450-272-8
16	49	3.8	7175	4	US-08-450-273-8
17	49	3.8	7177	4	US-09-268-163-7
18	49	3.8	7266	3	US-08-713-118-1
19	49	3.8	7266	3	US-09-452-007-1
20	49	3.8	7362	1	US-08-455-543A-7
21	49	3.8	7362	2	US-08-133-078B-7
22	49	3.8	7362	2	US-08-223-305C-7
23	49	3.8	7362	2	US-08-149-097D-7
24	49	3.8	7362	3	US-08-949-386-7
25	49	3.8	7362	3	US-08-450-562-7
26	49	3.8	7362	4	US-08-984-709A-7
27	49	3.8	7362	4	US-08-450-272-7

Sequence 7, Appli  
Sequence 5, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 51, Appl  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 163, App  
Sequence 159, App  
Sequence 163, App

#### ALIGNMENTS

RESULT 1  
US-07-745-206A-12  
; Sequence 12, Application US/07745206A  
; Patent No. 5429921  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: McCue, Ann  
; APPLICANT: Feldman, Daniel  
; TITLE OF INVENTION: Human Calcium Channel Compositions and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fitch, Even, Tabin & Flannery  
; STREET: 135 S. LaSalle  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/745,206A  
; FILING DATE: 19910815  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feder, Scott B  
; REFERENCE/DOCKET NUMBER: 51504  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-372-7842  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5467 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390  
; LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585  
; LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744  
; LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010  
; LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)  
US-07-745-206A-12











Mon Sep 27 09:07:51 2004

us-09-830-810a-1.rn1

```
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGGTGCAGGTGAACCCGCGCGGACGCTCGGTGCAGTGTCTACTCGGGCC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGGAGCCCGGTGCGCGGAGGAGCGGCCCGCGGCAC 2810
QY 310 GCGAGCTGCGAGCTGCGAGGTGCGGAGCGAGCCCGGACGCGCGGATCGGTCTCTGTC 369
Db 2811 GCGA-GCCACAGCAAGGAGCGCGCGGGCC--CCCGAGGCGCGGAGCGAGCGCGCGCA 2867
QY 370 CCCCGTGCACGCGCGCGCGGAGATCCCGCGCATCTGCGACACCGTAGCCCGCTTC 429
Db 2868 GCGCCAGGCGCGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2987
QY 490 CCCACAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 549
Db 2988 GCGCCAAAGGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3047
QY 550 GTGGCGCGGAGGAAAGCGGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 609
Db 3048 GCGGAGAGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3107
QY 610 GGGGAGCGCGGTGGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 668
Db 3108 CACGAGGCTGTGAGAAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3166
```

RESULT 9

US-08-193-078B-8  
; Sequence 8, Application US/08193078B

; Patent No. 5846757

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

; APPLICANT: Brenner, Robert

; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN

; STREET: 1660 UNION STREET

; CITY: SAN DIEGO

; STATE: CA

; COUNTRY: USA

; ZIP: 92101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/193,078B

; FILING DATE: 07-FEB-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/868,354

; FILING DATE: 10-APR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/745,206

; FILING DATE: 15-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie L.

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 6362-53607

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-238-0999

; TELEFAX: 619-238-0062

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7175 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 144..6857

; FEATURE:

; NAME/KEY: 5'UTR

; LOCATION: 1..143

; FEATURE:

; NAME/KEY: 3'UTR

; LOCATION: 6855..7175

; US-08-193-078B-8

Query Match 3.8%; Score 49; DB 2; Length 7175;

Best Local Similarity 48.0%; Pred. No. 0.0085;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGCAGGTGAACCCGCGCGGACGCTCGGTGCAGTGTCTACTCGGGCC 309

Db 2751 GAGGCCCGAAGCGGAGAGCGGGGAGCCCGGTGCGCGGAGGAGCGGCCCGCGGCAC 2810

QY 310 GCGAGCTGCGAGCTGCGAGGTGCGGAGCGAGCCCGGACGCGCGGATCGGTCTCTGTC 369

Db 2811 GCGA-GCCACAGCAAGGAGCGCGGGCC--CCCGAGGCGCGGAGCGGCGGCGGCGG 2867

QY 370 CCCCGTGCACGCGCGCGCGGAGATCCCGCGCATCTGCGACACCGTAGCCCGCTTC 429

Db 2868 GCGCCAGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2927

QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489

Db 2928 GAGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2987

QY 490 CCCACAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 549

Db 2988 GCGGCCAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3047

QY 550 GTGGCGCGGAGGAAAGCGGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 609

Db 3048 GCGGAGAGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3107

QY 610 GGGGAGCGCGGTGGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 668

Db 3108 CACGAGGCTGTGAGAAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3166

RESULT 10

US-08-223-305C-8

; Sequence 8, Application US/08223305C

; Patent No. 5851824

; GENERAL INFORMATION:

; APPLICANT: Harpold, Michael

; APPLICANT: Ellis, Steven

; APPLICANT: Williams, Mark

; APPLICANT: Feldman, Daniel

; APPLICANT: McCue, Ann

; APPLICANT: Brenner, Robert

; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brown, Martin, Haller & McClain

; STREET: 1660 Union Street

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101-2926



```

; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-55038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
; US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTCGGGTGCAGGTGAACCCGCGCGGAGACCGCTCGGTGAGTTTCACTCGGGCGC 309
DB 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGGTGCCCCGGAGGAGCGCGCGCGCCGAC 2810

QY 310 CGCAGGTGCAGCTGCAGGTGCCGAGCAGCCCGAGCCCGCATCGGTTCTCTCTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGGCGCGGGGCC--CCCGAGGCGCGGAGCGGAGCGCGCCG 2867

QY 370 CCCGTTGGCCACGCCGCGCGCGGAGATCCCCGCGATCTCTGGCAGACCGTAGCCCCGTT 429
DB 2868 GGCCCAAGCCCCAGGGCGCGCGGCGCACCA CGGGCGGCTCCCCGGAGGAGCGCGCC 2927

QY 430 TCGTCCGTGACCTTCTGTGGCCCTTCTCTCTCACTGGAGTTGGGGAGCGCAGCAGCA 489
DB 2928 GAGCGGAGCCCCGACGCCACCGCGCACCCGACCCAGGATCCGAGCAAGGAGTGCGCC 2987

QY 490 CCCACGAGGAGGAGGGAGCCCGGCATCTCTCGGGACCCCGGAACCGGAGCCGAGAG 549
DB 2988 GGGGCCAAGGGCGAGCGCGCGCGCGGACCGCGCGCGCCGCCGAGCGGGCCCCGGGAG 3047

QY 550 GTGGCCCGGAGAAAGCGGTCCCCCAGCGCGAAGCAGGAGGGCGGATGTTCAAGGTCGA 609
DB 3048 GCGAGAGCGGGAGAGCCCGCGGGCGGACCCGGCCCCGGCACAGGCGAGCCTGCT 3107

QY 610 GGGCAGGCCGGGTGGGAGCAGCAGCCACCCGAGGAGCCGGAACTGTGGGGCGCAT 668
DB 3108 CACGAGGCTGTGGAGAGGAGACCA CGGAGAAGGAGGCCACCGGAGAAGGAGGCTGAGAT 3166

RESULT 12
US-08-949-386-8
; Sequence 8, Application US/08949386
; Patent No. 6090623
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:

```



3166

```

RESULT 14
US-08-984-709A-8
; Sequence 8, Application US/08984709A
; Patent No. 6320032
;
GENERAL INFORMATION:
; APPLICANT: Williams, Mark E.
; APPLICANT: Straudman, Kenneth A.
; APPLICANT: Harpold, Michael M.
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, Suite 700
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,709A
; FILING DATE: 02-DEC-1997
; CLASSIFICATION: 435
;
ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
;
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 450-8400
; TELEFAX: (619) 587-5360
;
INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
;
FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
;
FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
;
US-08-984-709A-8

```

```
/ FILING DATE: 5-NOV-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/105,536
/ FILING DATE: 11-AUG-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/914,231
/ FILING DATE: 13-JULY-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/868,354
/ FILING DATE: 10-APR-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US92/06903
/ FILING DATE: 14-AUG-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/745,206
/ FILING DATE: 15-AUG-1991
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Seidman, Stephanie L.
/ REGISTRATION NUMBER: 33,779
/ REFERENCE/DOCKET NUMBER: 6362-519812
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 238-0999
/ TELEFAX: (619) 238-0062
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7175 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 144..6857
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: 1..143
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: 6855..7175
US-08-450-272-8

Query Match          3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0085;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTCGGTGACAGTGAAACCCCGCGCGACGCTCGGTGTCAGTGTTCACCTCGGGCGC 309
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2751 GAGGCCCGGAAGCGGAGAGCGGGAGCCCGGTGTCCTCGGGAGGAGCGCGCGCCGAC 2810

QY 310 CGCACGCTGCGAGCTGCGAGGGTGCGGAGCGAGCCCGACCGCCGATCGGTTCTGTCAA 369
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2811 CGCA-GCCAAGCAAGAGAGCCCGGGGCC--CCCGAGGGCGGAGCGAGCGCGCGCA 2867

QY 370 CCCCGTGGCCACCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGCCCGTTC 429
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2868 GGCACAGGCCCGAGGGCGCGCGGCAACCGGGCGGCTCCCGGAGGAGGCGGCC 2927

QY 430 TCCTCGTGACCTTCTGTGGCTCTCTCTCTACTGGAGGTGCGGAGGCGAGCAGACA 489
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2928 GAGCGGAGCGCCGACCGCCACCGCGGCACCGGCACCCAGATCCGAGCAGGAGTGGGCC 2987

QY 490 CCCACGAGGAGAGGGAGCCCGGATCTCTGGGACCCGGGAAACCGGAGCCGAGAGAG 549
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2988 GCGCCCAAGGGCGAGCGCGCGGCAACCGGGGCCCCCGGAGCGGGCCCCCGGAG 3047

QY 550 GTGGCCCGCAGGAAAGCGGTCCCCCAGCGCGAAGCGAGAGGGCGATGTTAGGGCTGCA 609
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3048 GCGGAGCGGGAGAGCGCGGCGCGCGCACCGGGCCCGGCACAAGCGCAGCCTGCT 3107
```

```
QY 610 GGCACGCGCGGTGGGAGCAGCAGCCACCGGAGGACCGGAACAGTGTGTGGCGCGAT 668
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3108 CACGAGGCTGTGGAGAAGGAGACCCAGGAGAGGAGCCACCGGAAGGAGGCTGAGAT 3166
```

Search completed: September 24, 2004, 20:03:21  
Job time : 121 secs

Blank